

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-125109

(43)Date of publication of application : 25.04.2003

(51)Int.Cl.

H04M 11/06  
G06F 3/16

(21)Application number : 2001-321045

(71)Applicant : HITACHI SOFTWARE ENG CO LTD

(22)Date of filing : 18.10.2001

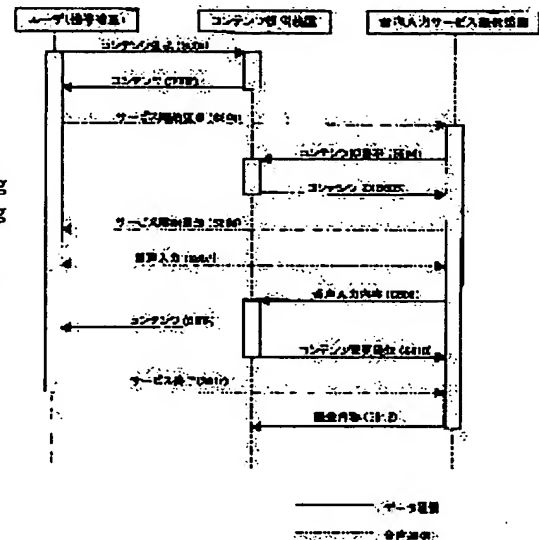
(72)Inventor : KOBAYASHI HIROYUKI

## (54) METHOD AND SYSTEM FOR PROVIDING VOICE INPUT SERVICE

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To provide a system and method for performing voice input to a contents information entry column displayed on a terminal synchronously with display of contents.

**SOLUTION:** The method of this invention includes; a first step of receiving a contents distribution request from the terminal via a data communication path by a contents providing device, discriminating whether or not the contents on request are contents set able to be in interlocking with the voice input serviced, attaching link destination information leading to a voice input service providing device to the contents according to the discriminated result, and distributing the resulting contents to the terminal being a requester; a second step of establishing the voice communication path between the concerned terminal and the voice input service providing device according to the selective operation of the distributed link destination information by the terminal; and a third step of recognizing the entered voice from the terminal through the voice communication path by the voice input service providing device and transmitting the result of recognition to the contents providing device.



## LEGAL STATUS

[Date of request for examination]

16.06.2004

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-125109

(P2003-125109A)

(43) 公開日 平成15年4月25日 (2003. 4. 25)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	ターミナル* (参考)
H 0 4 M 11/06		H 0 4 M 11/06	5 K 1 0 1
G 0 6 F 3/16	3 1 0	G 0 6 F 3/16	3 1 0 Z

審査請求 未請求 請求項の数13 O L (全 10 頁)

(21) 出願番号 特願2001-321045 (P2001-321045)

(22) 出願日 平成13年10月18日 (2001. 10. 18)

(71) 出願人 000233055

日立ソフトウェアエンジニアリング株式会社  
神奈川県横浜市鶴見区末広町一丁目1番43

(72) 発明者 小林 宏至

神奈川県横浜市中区尾上町6丁目81番地  
日立ソフトウェアエンジニアリング株式会  
社内

(74) 代理人 100088720

弁理士 小川 眞一

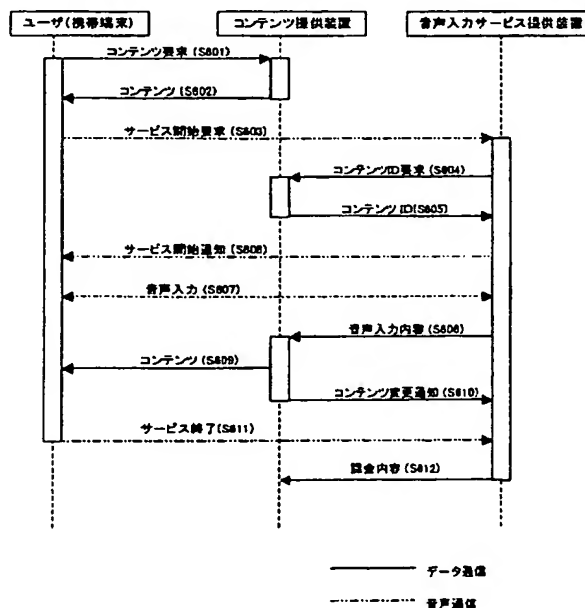
Fターム(参考) 5K101 KK02 KK15 KK16 KK17 LL00  
MM07 NN07 NN08 NN18 NN21  
PP03 VV07

## (54) 【発明の名称】 音声入力サービス提供方法及びシステム

## (57) 【要約】

【課題】 端末に表示されたコンテンツの情報入力欄に対する音声入力をコンテンツ表示と同期して実施可能にすること。

【解決手段】 端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテンツが音声入力サービスと連携可能に設定されたコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第1のステップと、配信した端末における前記リンク先情報の選択操作によって当該端末と音声入力サービス提供装置との音声通信路を確立する第2のステップと、音声入力サービス提供装置において音声通信路を介した端末からの入力音声を認識し、その認識結果を前記コンテンツ提供装置に送信する第3のステップとを備える。



**【特許請求の範囲】**

**【請求項 1】** データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援する方法であって、  
端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテンツが音声入力サービスと連携可能に設定されたコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第 1 のステップと、  
配信した端末における前記リンク先情報の選択操作によって当該端末と音声入力サービス提供装置との音声通信路を確立する第 2 のステップと、  
前記音声入力サービス提供装置において音声通信路を介した端末からの入力音声を認識し、その認識結果を前記コンテンツ提供装置に送信する第 3 のステップとを備えることを特徴とする音声入力サービス提供方法。

**【請求項 2】** データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援する方法であって、  
端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテンツが連携対象のコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第 1 のステップと、  
配信した端末における前記リンク先情報の選択操作によって当該端末と音声入力サービス提供装置との音声通信路を確立する第 2 のステップと、  
前記音声入力サービス提供装置において音声通信路を介した端末からの入力音声を認識し、その認識結果を前記コンテンツ提供装置に送信する第 3 のステップとを備えることを特徴とする音声入力サービス提供方法。

**【請求項 3】** データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援する方法であって、  
端末からのコンテンツ配信要求をコンテンツ提供装置で受付け、要求されたコンテンツ内の情報入力欄の有無に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第 1 のステップと、配信した端末における前記リンク先情報の選択操作によって当該端末から発信される音声入力サービス要求を音声入力サービス提供装置で受付け、サービス要求対象のコンテンツ識別情報を前記コンテンツ提供装置から取得する第 2 のステップと、  
取得したコンテンツ識別情報に対応するコンテンツに対する音声入力を支援するための音声入力支援情報を支援

情報データベースから取得し、当該音声入力支援情報をサービス要求元の端末に送信し、音声入力を指示する第 3 のステップと、

音声入力指示に従ってサービス要求元の端末で入力された音声を音声通信路を介して音声入力サービス提供装置で受信し、該入力音声を認識し、その認識結果を前記情報入力欄に対する入力情報として前記コンテンツ提供装置に送信する第 4 のステップとを備えることを特徴とする音声入力サービス提供方法。

10 **【請求項 4】** 前記第 1 のステップが、コンテンツ提供装置情報と、端末とコンテンツ提供装置とのセッション識別情報をリンク先情報中に付加して送信するステップを含み、  
前記第 2 のステップが、前記セッション識別情報をコンテンツ提供装置に送信し、当該セッション識別情報に対応するコンテンツ識別情報を取得するステップを含むことを特徴とする請求項 3 に記載の音声入力サービス提供方法。

20 **【請求項 5】** 音声入力サービス要求に従って実施した音声入力サービスに対する利用料金をコンテンツ提供装置の管理者に課金する第 5 のステップをさらに備えることを特徴とする請求項 3 または 4 に記載の音声入力サービス提供方法。

**【請求項 6】** コンテンツ提供装置と音声入力サービス提供装置とを備え、データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援するシステムであって、  
前記コンテンツ提供装置が、端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテンツが音声入力サービスと連携可能に設定されたコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第 1 の手段を備え、  
前記音声入力サービス提供装置が、  
30 配信した端末における前記リンク先情報の選択操作によって当該端末との音声通信路を確立する第 2 の手段と、  
音声通信路を介した端末からの入力音声を認識し、その認識結果を前記コンテンツ提供装置に送信する第 3 の手段とを備えることを特徴とする音声入力サービスシステム。

40 **【請求項 7】** コンテンツ提供装置と音声入力サービス提供装置とを備え、データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援するシステムであって、  
前記コンテンツ提供装置が、端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテンツが連携対象のコンテンツか否

かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第 1 の手段を備え、前記音声入力サービス提供装置が、配信した端末における前記リンク先情報の選択操作によって当該端末との音声通信路を確立する第 2 の手段と、音声通信路を介した端末からの入力音声認識し、その認識結果を前記コンテンツ提供装置に送信する第 3 の手段とを備えることを特徴とする音声入力サービスシステム。

【請求項 8】 コンテンツ提供装置と音声入力サービス提供装置とを備え、音声通信機能を有する端末で閲覧しているコンテンツの情報入力欄に対して音声による情報入力を支援するシステムであって、前記コンテンツ提供装置が、端末からのコンテンツ配信要求を受け、要求されたコンテンツ内の情報入力欄の有無に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第 1 の手段を備え、前記音声入力サービス提供装置が、コンテンツを配信した端末における前記リンク先情報の選択操作によって当該端末から発信される音声入力サービス要求を受け、サービス要求対象のコンテンツ識別情報を前記コンテンツ提供装置から取得する第 2 の手段と、取得したコンテンツ識別情報に対応するコンテンツに対する音声入力を支援するための音声入力支援情報を支援情報データベースから取得し、当該音声入力支援情報をサービス要求元の端末に送信し、音声入力を指示する第 3 の手段と、音声入力指示に従ってサービス要求元の端末で入力された音声を音声通信路を介して受信し、該入力音声を認識し、その認識結果を前記情報入力欄に対する入力情報として前記コンテンツ提供装置に送信する第 4 の手段とを備えることを特徴とする音声入力サービス提供システム。

【請求項 9】 前記第 1 の手段が、コンテンツ提供装置情報と、端末とコンテンツ提供装置とのセッション識別情報をリンク先情報中に付加して送信する手段を含み、前記第 2 の手段が、前記セッション識別情報をコンテンツ提供装置に送信し、当該セッション識別情報に対応するコンテンツ識別情報を取得する手段を含むことを特徴とする請求項 8 に記載の音声入力サービス提供システム。

【請求項 10】 前記音声入力サービス提供装置が、音声入力サービス要求に従って実施した音声入力サービスに対する利用料金をコンテンツ提供装置の管理者に課金する第 5 の手段をさらに備えることを特徴とする請求項 8 または 9 に記載の音声入力サービス提供システム。

【請求項 11】 前記コンテンツ提供装置が、端末に配信するコンテンツを格納したコンテンツデータベースと、格納されたコンテンツの情報入力欄に対する音声入

力支援情報を作成する支援情報作成装置とを備え、前記支援情報作成装置が、配信対象のコンテンツ内における情報入力欄を抽出する手段と、抽出した情報入力欄に対する音声入力支援用の情報を設定し、設定結果の音声入力支援情報を前記音声入力サービス提供装置に送信し、前記支援情報データベースに格納させる手段とを備えることを特徴とする請求項 8 ～ 10 のいずれか一項に記載の音声入力サービス提供システム。

【請求項 12】 前記音声入力サービス提供装置が、前記支援情報データベースから取得した音声入力支援情報に対応する合成音声を生成あるいは録音された音声を再生し、サービス要求元の端末に音声通話路を介して送信する手段を備えることを特徴とする請求項 8 ～ 11 のいずれか一項に記載の音声入力サービス提供システム。

【請求項 13】 パスワードなどの音声入力に適さない入力欄に対しては、前記音声入力支援情報の設定を禁止する手段をさらに備えることを特徴とする請求項 11 に記載の音声入力サービス提供システム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、携帯電話機などの端末に表示したコンテンツに対して、音声によって必要情報を入力するサービスを提供する音声入力サービス提供方法およびシステムに関するものである。

【0002】

【従来の技術】 最近のインターネットに接続可能な携帯電話機の普及に伴い、携帯電話機利用者を対象とした電子商取引のコンテンツが提供されるようになった。このようなコンテンツではしばしば、商品検索やユーザ登録など利用者が文字を入力する必要がある。従来、携帯電話機でコンテンツに文字を入力するには、携帯電話機に付随する主に 0 ～ 9 のキーを使用していた。

【0003】

【発明が解決しようとする課題】 しかしながら、利用者が 10 個程度のキーで英数字、ひらがな、カタカナ、漢字を入力するのは難しく、このことが携帯電話機を利用した電子商取引の普及を妨げている。そこで、キー入力に代わる文字の入力方法として、音声入力方法が考えられている。これには、音声認識及び合成を行う音声入力システムを携帯電話機側に置く方法とコンテンツの配信サーバ側に置く方法が考えられる。前者の場合、携帯電話機の処理能力の問題から、音声認識能力の高い音声入力システムは実現できない。後者の場合、これまで音声入力のための音声通信とコンテンツを通信するデータ通信が同時に行えなかったため実現できなかった。

【0004】 次世代移動通信では音声通信とデータ通信を同時に行うことができるが、このとき問題になるのが、音声通信とデータ通信の同期を取ることである。すなわち、音声通信とデータ通信はお互いに別々のチャンネルで通信される。そのため、音声通信を処理している

10

30

40

50

システムは利用者が現在閲覧しているコンテンツを認識できない。また、データ通信を処理しているシステムは、音声入力されたデータがどの時点で入力されたか特定できず、どのページのどの入力欄に対応するものか把握できない。

【0005】本発明の目的は、このような課題を解決し、音声とデータの同時通信サービスを利用し、携帯電話機等の端末に表示されたコンテンツに対する音声入力をコンテンツ表示と同期して実施することができる音声入力サービス提供方法及びシステムを提供することにある。

【0006】

【課題を解決するための手段】上記の目的を達成するために、本発明は、データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援する方法であって、端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテンツが音声入力サービスと連携可能に設定されたコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第1のステップと、配信した端末における前記リンク先情報の選択操作によって当該端末と音声入力サービス提供装置との音声通信路を確立する第2のステップと、前記音声入力サービス提供装置において音声通信路を介した端末からの入力音声を認識し、その認識結果を前記コンテンツ提供装置に送信する第3のステップとを備えることを特徴とする。

【0007】また、データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援する方法であって、端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテンツが連携対象のコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第1のステップと、配信した端末における前記リンク先情報の選択操作によって当該端末と音声入力サービス提供装置との音声通信路を確立する第2のステップと、前記音声入力サービス提供装置において音声通信路を介した端末からの入力音声を認識し、その認識結果を前記コンテンツ提供装置に送信する第3のステップとを備えることを特徴とする。

【0008】また、データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援する方法であって、端末からのコンテンツ配信要求をコンテンツ提供装置で受付け、要求されたコンテンツ内の情報入力欄の有無に応じて、当該コンテンツに音声入

力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第1のステップと、配信した端末における前記リンク先情報の選択操作によって当該端末から発信される音声入力サービス要求を音声入力サービス提供装置で受付け、サービス要求対象のコンテンツ識別情報を前記コンテンツ提供装置から取得する第2のステップと、取得したコンテンツ識別情報に対応するコンテンツに対する音声入力を支援するための音声入力支援情報を支援情報データベースから取得し、当該音声入力支援情報をサービス要求元の端末に送信し、音声入力を指示する第3のステップと、音声入力指示に従ってサービス要求元の端末で入力された音声を音声通信路を介して音声入力サービス提供装置で受信し、該入力音声を認識し、その認識結果を前記情報入力欄に対する入力情報として前記コンテンツ提供装置に送信する第4のステップとを備えることを特徴とする。また、前記第1のステップが、コンテンツ提供装置情報と、端末とコンテンツ提供装置とのセッション識別情報をリンク先情報中に付加して送信するステップを含み、前記第2のステップが、前記セッション識別情報をコンテンツ提供装置に送信し、当該セッション識別情報に対応するコンテンツ識別情報を取得するステップを含むことを特徴とする。また、音声入力サービス要求に従って実施した音声入力サービスに対する利用料金をコンテンツ提供装置の管理者に課金する第5のステップをさらに備えることを特徴とする。

【0009】また、コンテンツ提供装置と音声入力サービス提供装置とを備え、データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援するシステムであって、前記コンテンツ提供装置が、端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテンツが音声入力サービスと連携可能に設定されたコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第1の手段を備え、前記音声入力サービス提供装置が、配信した端末における前記リンク先情報の選択操作によって当該端末との音声通信路を確立する第2の手段と、音声通信路を介した端末からの入力音声を認識し、その認識結果を前記コンテンツ提供装置に送信する第3の手段とを備えることを特徴とする。

【0010】また、コンテンツ提供装置と音声入力サービス提供装置とを備え、データ通信機能と音声通信機能を有する端末でネットワークコンテンツの閲覧中に、当該ネットワークコンテンツに対して音声による情報入力を支援するシステムであって、前記コンテンツ提供装置が、端末からのコンテンツ配信要求をデータ通信路を介してコンテンツ提供装置で受付け、要求されたコンテン

10

20

30

40

50

ツが連携対象のコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第1の手段を備え、前記音声入力サービス提供装置が、配信した端末における前記リンク先情報の選択操作によって当該端末との音声通信路を確立する第2の手段と、音声通信路を介した端末からの入力音声を認識し、その認識結果を前記コンテンツ提供装置に送信する第3の手段とを備えることを特徴とする。

【0011】また、コンテンツ提供装置と音声入力サービス提供装置とを備え、音声通信機能を有する端末で閲覧しているコンテンツの情報入力欄に対して音声による情報入力を支援するシステムであって、前記コンテンツ提供装置が、端末からのコンテンツ配信要求を受け、要求されたコンテンツ内の情報入力欄の有無に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信する第1の手段を備え、前記音声入力サービス提供装置が、コンテンツを配信した端末における前記リンク先情報の選択操作によって当該端末から発信される音声入力サービス要求を受け、サービス要求対象のコンテンツ識別情報を前記コンテンツ提供装置から取得する第2の手段と、取得したコンテンツ識別情報に対応するコンテンツに対する音声入力を支援するための音声入力支援情報を支援情報データベースから取得し、当該音声入力支援情報をサービス要求元の端末に送信し、音声入力を指示する第3の手段と、音声入力指示に従ってサービス要求元の端末で入力された音声を音声通信路を介して受信し、該入力音声を認識し、その認識結果を前記情報入力欄に対する入力情報として前記コンテンツ提供装置に送信する第4の手段とを備えることを特徴とする。また、前記第1の手段が、コンテンツ提供装置情報と、端末とコンテンツ提供装置とのセッション識別情報をリンク先情報中に付加して送信する手段を含み、前記第2の手段が、前記セッション識別情報をコンテンツ提供装置に送信し、当該セッション識別情報に対応するコンテンツ識別情報を取得する手段を含むことを特徴とする。また、前記音声入力サービス提供装置が、音声入力サービス要求に従って実施した音声入力サービスに対する利用料金をコンテンツ提供装置の管理者に課金する第5の手段をさらに備えることを特徴とする。また、前記コンテンツ提供装置が、端末に配信するコンテンツを格納したコンテンツデータベースと、格納されたコンテンツの情報入力欄に対する音声入力支援情報を作成する支援情報作成装置とを備え、前記支援情報作成装置が、配信対象のコンテンツ内における情報入力欄を抽出する手段と、抽出した情報入力欄に対する音声入力支援用の情報を設定し、設定結果の音声入力支援情報を前記音声入力サービス提供装置に送信し、前記支援情報データベースに格納させる手段とを備えることを特徴とする。また、前記音声入力サービス提

供装置が、前記支援情報データベースから取得した音声入力支援情報に対応する合成音声を生成あるいは録音された音声を再生し、サービス要求元の端末に音声通話路を介して送信する手段を備えることを特徴とする。また、パスワードなどの音声入力に適さない入力欄に対しては、前記音声入力支援情報の設定を禁止する手段をさらに備えることを特徴とする。

#### 【0012】

【発明の実施の形態】以下、本発明を実施する場合の一形態を、図面を参照して具体的に説明する。なお、本発明は、携帯電話機に限らず、音声通信機能を有するPDA (Personal Digital (Data) Assistants) などの携帯情報端末であっても実現可能であるため、以下では、携帯電話機やPDAを総称して携帯端末と言う。また、データ通信と音声通信とを同時に行うことができる構成であっても実施することができる。例えば、ネットワークとのデータ通信機能を備えたデスクトップコンピュータやノート型コンピュータ、あるいはインターネット接続機能を備えたテレビジョンに対し、固定電話網またはインターネット電話網を介して音声通信を行う機能を付加した構成であっても実施することができる。また、インターネット接続機能を備えたファクシミリ装置であっても送受話器が付属しているものであれば、適用することができる。以下の実施形態では、データ通信機能と音声通信機能を有する携帯電話機に適用した例を詳しく説明する。

【0013】図1は、本発明に係る音声入力サービス提供システムの実施の形態を示す全体構成図である。ここで示すシステムは、ユーザが情報入力欄を有するコンテンツを閲覧し、情報入力欄に対して音声によって必要情報を入力する携帯端末110、この携帯端末110の間では音声とデータとを同時に無線通信し、インターネット140を通してデータを、固定電話網150を通して音声を通信する基地局120、インターネット140と無線通信網との間のプロトコルの相互変換を行うゲートウェイ130、データを通信するインターネット140、音声を通信する固定電話網150、携帯端末110にコンテンツを提供するコンテンツ提供装置160、このコンテンツ提供装置160に音声入力サービスを提供し、コンテンツ提供装置160が提供しているコンテンツに対してユーザによる音声入力を可能にする音声入力サービス提供装置170から構成される。

【0014】コンテンツ提供装置160は、コンテンツを保持するコンテンツデータベース(DB)162、このコンテンツDB162からコンテンツを取得し、携帯端末110にコンテンツを配信するWebサーバ161、コンテンツに対する音声入力に必要な支援情報(音声入力を促すための問い合わせ音声あるいは案内音声を生成するのに必要な情報)を作成する音声入力支援情報作成パーソナルコンピュータ(PC)163から構成さ

れる。

【0015】音声入力サービス提供装置170は、固定電話網150のアナログ音声デジタル情報に変換すると共に逆の変換を行うPBX171、音声入力に必要な支援情報を保持する音声入力支援情報データベース(DB)173、この音声入力支援情報DB173が保持する支援情報に基づきユーザに音声入力を促すための案内音声または問い合わせ音声を合成あるいは再生して携帯端末110に配信すると共に、ユーザが入力した音声を認識する音声認識・合成サーバ172から構成される。

【0016】このようなシステム構成において、コンテンツ提供装置160が音声入力サービス提供装置170と契約し、コンテンツ提供装置160が配信しているコンテンツに対し、音声入力サービスをユーザに提供する手順を図2に示す。まず、ステップS201では、コンテンツ提供装置160は音声入力サービス提供装置170と音声入力サービスに関して契約を行う。このとき、音声入力サービス提供装置170は、コンテンツ提供装置160に音声入力に必要な情報(音声入力支援情報)を作成するための音声入力支援情報作成プログラムを与える。

【0017】ステップS202では、コンテンツ提供装置160は、与えられた音声入力支援情報作成プログラムを音声入力支援情報作成PC163上で実行し、音声入力を行うのに必要な案内音声などの音声入力支援情報を作成、音声入力サービス提供装置170に送信する。ステップS203では、コンテンツ提供装置160は、コンテンツに対する音声入力サービスを実施する。ステップS204では、音声入力サービス提供装置170は、サービスの利用状況に応じたサービス提供料金をコンテンツ提供装置160の管理者に課金(請求)する。

【0018】図3は、ステップS202の音声入力支援情報作成処理工程における音声入力支援情報作成PC163の画面構成例を示す図である。これは、音声入力サービスの対象となるコンテンツを表示するコンテンツ表示画面D310と、コンテンツにおける氏名、フリガナ、クレジットカード番号等の情報入力欄(図で矩形白枠で示す部分)に対して音声入力支援情報を設定する音声入力支援情報設定画面D320から構成される。音声入力支援情報設定画面D320には、コンテンツの各入力欄に対する設定を行うための入力設定部P321が表示されている。

【0019】図4にコンテンツ提供装置160が音声入力支援情報作成PC163を用いて音声入力支援情報を作成・配信するフローチャートを示す。まず、ステップS401では、URLを指定してWebサーバ161から音声入力支援情報の作成対象となる情報入力欄を含むコンテンツを取得し、コンテンツ表示画面D310に表示する。ステップS402では、コンテンツのタグ<FORM>、<INPUT>、<TEXTAREA>、<

SELECT>から氏名等の情報入力欄を判断し、各入力欄に対して音声入力を許可するかどうか等の設定を行う入力設定部P321を音声入力支援情報設定画面D320に表示する。

【0020】ステップS403では、設定を行う情報入力欄に対応する1つの入力設定部P321を選択する。入力設定部P321は、有効、項目、読み、入力形式、録音の各項目から構成され、音声入力を許可する場合には「有効」の項目を「×印」に設定し、許可しない場合には、空欄とする。ステップS404では、その情報入力欄に対して音声入力を許可するか選択する。図3の例では他人に聞かれると問題となるクレジットカード番号は、音声入力を許可していない。不許可の場合はステップS403に戻り、別の情報入力欄に対応する入力設定部P321を選択する。許可の場合はステップS405を実行する。

【0021】ステップS405では、音声入力の際にユーザに入力欄に対する情報入力を音声で指示するときの案内音声または問い合わせ音声に対応する文を設定する。これには、2通りあり、「読み」の項目にローマ字で設定し、音声入力の際には音声認識・合成サーバ172で合成してユーザに指示する場合と、音声を録音して音声ファイルに記録しておき、音声入力の際に録音した音声を発音させる方法がある。ローマ字で設定するのは、音声合成をし易くするためである。図3の例では「生年月日」の項目に対し「録音」が設定されており、この「生年月日」の項目だけ録音し、それ以外はローマ字で設定された通りの音声を合成して発音させるように設定している。

【0022】ステップS406では、ユーザが入力した音声を音声認識・合成サーバ172が認識して文字に変換するときの入力形式を指定する。入力形式としては、全角漢字、カタカナ、英数字、数字などの種類がある。ステップS407では、入力設定部P321への音声入力支援情報の入力完了した場合は、ステップS408を実行し、そうでない場合はステップS403を実行し、別の情報入力欄に対応する入力設定部P321の設定に移る。

【0023】ステップS408では、画面上部のメニューの「送信」を選択し、作成した音声入力支援情報を音声入力サービス提供装置170者に送信する。この場合、コンテンツ提供装置170を特定するためのコンテンツ提供装置ID、音声入力サービスを実施するコンテンツを特定するコンテンツID、各情報入力欄に対応する案内音声の出力順序の情報を付加して送信する。

【0024】音声入力サービス提供装置170は、受信した音声入力支援情報を音声入力支援情報DB173に格納する。音声入力支援情報DB173が保持する音声入力支援情報は、図5に示すように、コンテンツ提供装置ID501、コンテンツID502、順序503、項

10

20

30

40

50



目504、読み505、入力形式506、音声ファイル507の項目で構成されている。ここで、音声ファイル507とは、録音音声を使用する場合の音声を記録した音声ファイルのファイル名称を指すものである。

【0025】次に、ステップS203の音声入力サービスを実施する場合の手順について、図6を参照して説明する。まず、ステップS601において、ユーザが携帯端末110からWebサーバ161にコンテンツを要求する。ステップS602では、Webサーバ161が要求されたコンテンツをコンテンツDB162から取得する。このとき、取得したコンテンツに氏名等の情報入力欄が含まれている場合は、Webサーバ161はユーザが音声入力サービスを利用するための音声入力サービス提供装置170へのリンクをコンテンツに付加する。なお、情報入力欄が含まれていても、音声入力サービス提供装置170とのサービス契約をしていないコンテンツ提供装置であった場合には、リンクを設定しない。

【0026】リンクを付加したコンテンツの例を図7に示す。図の左側がコンテンツをブラウザ表示したもので、図の右側がHTML形式で表示したものであり、付加した部分は点線701、702で囲んで示している。図の左のリンク先(属性cti)には、音声入力サービス提供装置170への電話番号「0123456」とトーン入力部にはコンテンツ提供装置160のID「12345」とユーザのセッションID「3.4.56」が指定されている。トーン入力部のID「12345」と「3456」は、音声入力サービス提供装置170との音声通話路が確立した後、これらのIDをプッシュボタン信号に相当する多周波信号で伝達するためのものである。

【0027】ステップS603において、ユーザが音声入力サービスを受けるため、付加された音声入力サービス提供装置170へのリンク701を選択すると、携帯端末110がHTMLデータのリンク72で指定された電話番号の音声入力サービス提供装置170への呼び出し信号を発信し、音声入力サービス提供装置170との通話路を確立し、音声による情報入力のサービス要求を行う。

【0028】音声通信による通話路が確立したならば、ステップS604において、PBX171が携帯端末110からのトーン信号入力により、ユーザが閲覧しているコンテンツ提供装置170とユーザのセッションIDを取得し、そのコンテンツ提供装置のWebサーバ161にセッションIDを送信し、ユーザが閲覧しているコンテンツIDを要求する。ステップS605では、Webサーバ161が、受信したセッションIDからユーザが閲覧しているコンテンツIDを音声認識・合成サーバ172に送信する。ステップS606では、音声認識・合成サーバ172は、コンテンツIDを受信すると音声入力サービスの準備が整い、サービスを開始することを音声でユーザに通知する。

【0029】ステップS607では、音声入力を実施する。このとき、音声認識・合成サーバ172が行う手順を図8のフローチャートを参照して説明する。まず、ステップS801では、コンテンツIDを基に音声入力支援情報DB173から当該コンテンツIDに対応するコンテンツ内の情報入力欄に対する音声入力支援情報を取得する。ステップS802では、取得した音声入力支援情報に問合せ用の「音声ファイル」が存在するか判定する。存在する場合はステップS803を実行し、存在しない場合はステップS804を実行する。図5の例では「誕生日」に関する問合せだけ音声ファイルが存在する。

【0030】ステップS803では、図5で指定された「音声ファイル」を再生してユーザに問合せる。例えば、「誕生日を音声で入力して下さい」といった問い合わせ音声を発音させ、誕生日の情報を入力させる。ステップS804では、図5で「読み」の項目の案内音声を音声合成し、ユーザに問合せる。例えば、「shimei」という読みに対し、「氏名を音声で入力して下さい」といった問い合わせ音声を発音させ、氏名の情報を入力させる。ステップS805では、ユーザからの音声による返答を認識する。

【0031】ステップS806では、認識した音声を図5の「入力形式」のフォーマットのデータに変換する。もし変換できない場合は、ステップS802を実行して再度ユーザに問合せる。変換できる場合は、ステップS807を実行する。ステップS807では、コンテンツ内の音声入力を許可した全ての情報入力欄の入力を完了したことを判定する。完了していなければステップS801を実行し、次の情報入力欄の問合せを行う。

【0032】図6に戻り、ステップS608では、音声認識・合成サーバ172は、入力された情報(音声入力内容)をユーザのセッションIDとともにコンテンツ提供装置160のWebサーバ161に送信する。ステップS609では、Webサーバ161は音声認識・合成サーバ172からの入力情報とセッションIDを受信し、入力情報に基づいて情報入力欄に入力情報を文字入力し、返答のコンテンツとして、セッションIDを基にユーザの携帯端末110に配信する。

【0033】ステップS610では、Webサーバ161はステップS609で携帯電話110に返答のコンテンツを送信したことを通知するために、音声認識・合成サーバ172へ新たなコンテンツのIDを送信する。また、携帯電話110からのコンテンツ要求により、Webサーバ161が携帯電話110に要求されたコンテンツを送信したときも、同様に音声認識・合成サーバ172へ新たなコンテンツIDを送信する。新たなコンテンツIDを受信した音声認識・合成サーバ172は、音声入力情報DB173を参照してそのコンテンツに音声入力可能な情報入力欄が含まれているか判定し、含まれて



いる場合はステップ S 6 0 7 を実行し、音声入力を行う。

【0034】ステップ S 6 1 1 では、音声入力を終了する場合、ユーザは音声入力サービス提供装置 1 7 0 との音声通話路を切断する。ステップ S 6 1 2 では、PBX 1 7 1 は、ユーザへのサービス提供が終了したことを判定し、ユーザのサービス利用時間、利用ユーザ数、入力ページ数、入力データ量等を Web サーバ 1 6 1 に通知する。後日、これらの情報をもとに音声入力サービス提供装置 1 7 0 はコンテンツ提供装置 1 6 0 の管理者に利用料金を請求する。

【0035】以上説明したように、本実施形態によれば、携帯端末のユーザが音声入力サービスの提供を音声入力サービス提供装置に音声通話路を介して要求し、その際に、閲覧しているコンテンツのセッション ID をトーン信号で音声入力サービス提供装置に伝達し、音声入力サービス提供装置にコンテンツ提供装置からセッション ID に対応するコンテンツ ID を取得させることによって、ユーザが現在閲覧しているコンテンツを認識させ、そのコンテンツの情報入力欄に対する音声入力の案内音声を表示内容と同期して発音させ、必要な情報を音声によって入力させることができる。また、コンテンツ提供者は音声入力支援情報作成 PC により、音声入力の許可、入力欄を指定するための問合せ文（案内音声）に関する音声入力支援情報を容易に作成することができる。

【0036】なお、上記実施形態では、本発明を携帯電話機に適用した例を挙げて説明したが、本発明はこれに限定されるものではなく、データ通信と音声通信とを同時に行うことができる構成であれば、同様に適用することができる。例えば、ネットワークとのデータ通信機能を備えたデスクトップコンピュータやノート型コンピュータ、あるいはインターネット接続機能を備えたテレビジョンに対し、固定電話網またはインターネット電話網を介して音声通信を行う機能を付加した構成であっても実施することができる。また、インターネット接続機能を備えたファクシミリ装置であっても送受信器が付属しているものであれば、適用することができる。

【0037】また、コンテンツ中の情報入力欄に音声入力する例を説明したが、情報入力欄の有無に関係なく、各コンテンツに対し音声入力と連携可能が否かのフラグを設定しておき、端末からの要求があった場合には、要求されたコンテンツが音声入力サービスと連携可能に設定されたコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信し、配信した端末における前記リンク先情報の選択操作によって当該端末と音声入力サービス提供装置との音声通信路を確立した後、音声入力サービス提供装置において音声通信路を介した端末からの入力音声を認識し、その認識結果

をコンテンツ提供装置に送信するという構成にすることができる。

【0038】また、同様に、要求されたコンテンツが音声入力と連携対象のコンテンツか否かを判定し、その判定結果に応じて、当該コンテンツに音声入力サービス提供装置へのリンク先情報を付加して要求元の端末に配信し、その配信した端末における前記リンク先情報の選択操作によって当該端末と音声入力サービス提供装置との音声通信路を確立するように構成することができる。これは、例えば、コンテンツ中に質問文のみを表示し、回答を音声で受けるといったアンケート等のコンテンツに適用することができる。アンケートコンテンツが、前記の音声入力サービスと連携可能に設定されたコンテンツまたは音声入力と連携対象のコンテンツに相当する。

【0039】

【発明の効果】以上の説明から明らかなように、本発明によれば、音声とデータの同時通信機能を利用し、携帯電話機等の端末に表示されたコンテンツの情報入力欄に対する音声入力をコンテンツ表示と同期して実施することができる。

【図面の簡単な説明】

【図 1】本発明の一実施形態を示すシステム構成図である。

【図 2】コンテンツ提供装置が音声入力サービス提供装置からサービスを受ける手順を示すフローチャートである。

【図 3】情報入力欄を含むコンテンツに対して音声入力支援情報を作成する音声入力支援情報作成 PC の画面例を示す図である。

【図 4】コンテンツ提供装置が音声入力支援情報作成 PC を用いて音声入力支援情報を作成する手順を示すフローチャートである。

【図 5】音声入力支援情報データベースが保持している音声入力支援情報の例を示す図である。

【図 6】音声入力サービスを実施するときのユーザとコンテンツ提供装置および音声入力サービス提供装置のやりとりを示すシーケンス図である。

【図 7】情報入力欄を含むコンテンツに音声入力サービスを開始するための音声入力サービス提供装置へのリンクを付加した例を示す図である。

【図 8】音声入力時の音声認識・合成サーバの動作を示すフローチャートである。

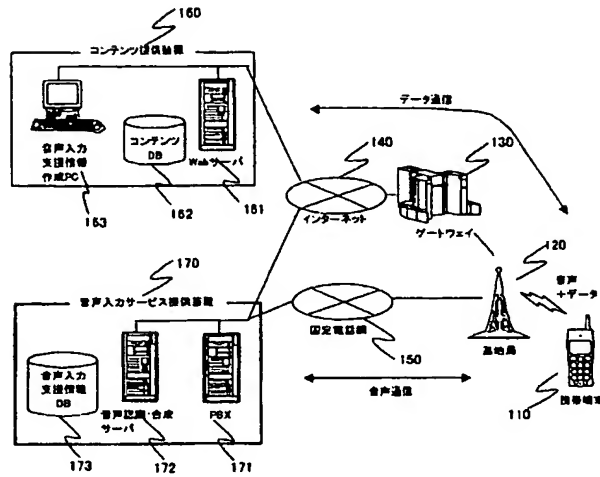
【符号の説明】

110…携帯端末、120…基地局、130…ゲートウェイ、140…インターネット、150…固定電話網、160…コンテンツ提供装置、161…Web サーバ、162…コンテンツデータベース、163…音声入力支援情報作成 PC、170…音声入力サービス提供装置、171…PBX、172…音声認識・合成サーバ、173…音声入力支援情報データベース、D310…コンテ

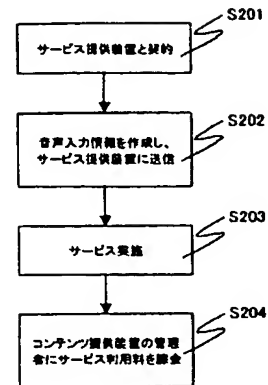
ンツ表示画面、D320…音声入力支援情報設定画面、

P321…入力設定部。

【図1】



【図2】



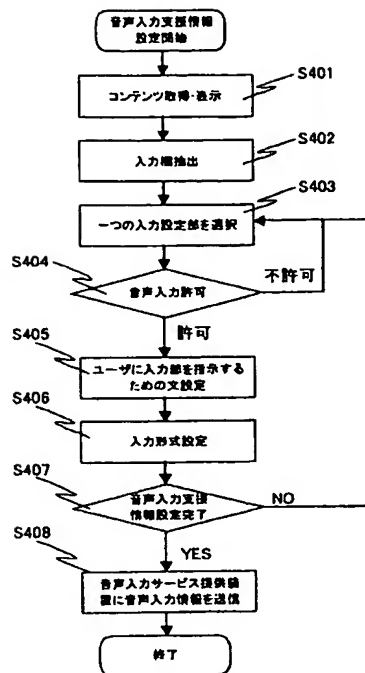
【図3】

Figure 3 shows two screens. The left screen (D310) is the 'ユーザ登録' (User Registration) screen with fields for: 氏名 (Name), フリガナ (Kana), 住所 (Address), 電話番号 (Phone number), E-mail, 生年月日 (Date of birth), and クレジットカード番号 (Credit card number). It has '送信' (Send) and 'キャンセル' (Cancel) buttons. The right screen (D320) is the '入力設定' (Input Setting) screen with a table for setting input methods.

有効	項目	読み	入力形式	録音
<input checked="" type="checkbox"/>	氏名	shimei	全角漢字	○
<input checked="" type="checkbox"/>	フリガナ	hurigana	カタカナ	○
<input checked="" type="checkbox"/>	住所	jyusho	全角漢字	○
<input checked="" type="checkbox"/>	電話番号	dennwabago	数字	○
<input checked="" type="checkbox"/>	E-mail	i-meiru	英数字	○
<input checked="" type="checkbox"/>	生年月日		数字	●
<input type="checkbox"/>	クレジットカード番号			○

Labels P321 and D320 point to the input setting table, and D310 points to the registration form.

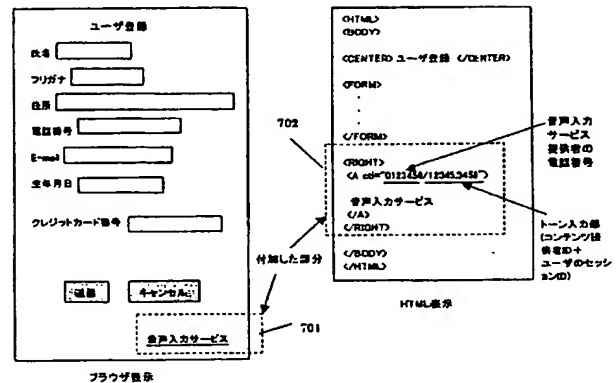
【図4】



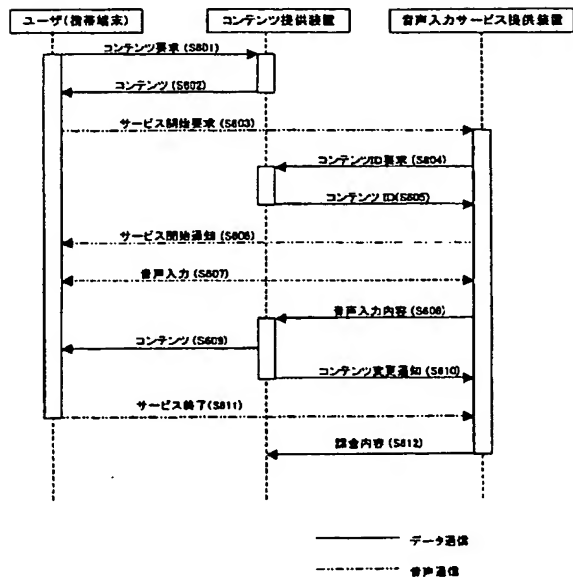
【図5】

コンテンツ提供装置	コンテンツID	順序	項目	読み	入力形式	音声ファイル
12345	WP001	1	氏名	shimai	全角漢字	NO
12345	WP001	2	フリガナ	hurigana	カタカナ	NO
12345	WP001	3	住所	kyusho	全角漢字	NO
12345	WP001	4	電話番号	denwabango	数字	NO
12345	WP001	5	E-mail	l-mail	英数字	NO
12345	WP001	6	生年月日	NO	数字	Birth.wav

【図7】



【図6】



【図8】

